

Leveraging over intact priors for boosting control and dexterity of prosthetic hands by amputees

Valentina Gregori and Barbara Caputo

Abstract—Non-invasive myoelectric prostheses require a long training time to obtain satisfactory control dexterity. These training times could possibly be reduced by leveraging over training efforts by previous subjects. So-called domain adaptation algorithms formalize this strategy and have indeed been shown to significantly reduce the amount of required training data for intact subjects for myoelectric movements classification. It is not clear, however, whether these results extend also to amputees and, if so, whether prior information from amputees and intact subjects is equally useful. To overcome this problem, we evaluated several domain adaptation algorithms on data coming from both amputees and intact subjects. Our findings indicate that: (1) the use of previous experience from other subjects allows us to reduce the training time by about an order of magnitude; (2) this improvement holds regardless of whether an amputee exploits previous information from other amputees or from intact subjects.

Index Terms—Prostheses, EMG signal, training time, domain adaptation, learning.

I. INTRODUCTION

A majority of upper-limb amputees is interested in prostheses controlled via surface electromyography (EMG), but they perceive the prohibitively difficult control as a great concern [1]. This becomes even more problematic with highly articulated modern prostheses. The natural use of these devices is challenging in everyday life primarily due to the software [2, 3, 4]. The open question is how to reduce this training time while making control as natural as possible. Machine learning has opened a new path to tackle this problem by allowing the prosthesis to adapt to the myoelectric signals of a specific user. Although these methods have been applied with success (e.g., [5] and references therein), they still require a significant amount of data from individual subjects to learn models with satisfactory performance.

Consider a situation in which different subjects repeat the same hand postures and suppose that a new target user attempts to learn the same movements. In this case it should be beneficial to reuse the information from the latter subjects and thereby reduce the training data required from a new subject. However, even if a movement appears the same for all subjects, the distribution of their myoelectric signals is very different. These depend on user characteristics like age, subcutaneous fat, muscular activity and on the position of the electrodes

during the data collection [6]. Domain adaptation methods can be used to overcome this mismatch between these different data distributions by considering each subject as an individual domain. This allows boosting prosthetic control by exploiting earlier training efforts by other source subjects.

These techniques have shown promising results for myoelectric movements classification with a large amount of intact subjects [7, 8]. However, these methods have never been verified on a significant number of amputees. Would the findings presented so far still be valid in these cases, or would the differences in the signal caused by the amputation be so strong as to prevent any useful knowledge transfer? And would it be possible to leverage over intact sources to boost control of amputees?

This paper aims at answering these questions. We assess if and how previous knowledge from intact subjects or amputees can reduce the time needed by an amputee to learn to control the device. To this end, we ran three experiments. In the first we considered only intact subjects. Then, we took into account amputated subjects that exploit previous knowledge either from other amputees or from intact subjects (second and third experiment). In order to clarify the importance of prior knowledge, we employed several existing adaptive learning algorithms using EMG data from the largest public database (NinaPro [9, 10, 11]).

The rest of the paper is organized as follows. In section II we summarize the main contributions of previous work. Subsequently we introduce the algorithms used in this work (section III), the data and their processing (section IV). In section V we present our results. Section VI is a conclusive summary that focuses on the main findings.

II. RELATED WORK

One of the first work that attempted to classify EMG signals of three volunteers is by Graupe and Cline [12]. In the following years studies on prosthetic control continued in various directions leading to many advances in the analysis and understanding of EMG. Few pioneer works highlight that EMG signals differ from subject to subject and that signals from different subjects are therefore not automatically reusable [13]. They proposed a cross-subject analysis to show that reusing the pre-trained models from former subjects can shorten the training time for a new user. Matsubara et al. [14] proposed to extract a user-independent component from EMG data that is considered transferable across subjects.

Different works applied several domain adaptation methods for the task of EMG recognition, showing that these algorithms

V. Gregori is with Department of Computer, Control, and Management Engineering, University of Rome La Sapienza, via Ariosto 25, 00185 Roma, Italy.

B. Caputo is with Department of Computer, Control, and Management Engineering, University of Rome La Sapienza, via Ariosto 25, 00185 Roma, Italy.

outperform non-adaptive ones [15, 16]. Chattopadhyay et al. [17] proposed an adaptation method for EMG classification based on a weighted combination of target and source samples. The shortcoming in these studies is that they involved a small number of subjects. Only recent studies [7, 8], done after the release of the NinaPro database [9], took into account a significant amount of intact subjects (from 10 to 40) and postures (from 12 to 52). They tested several domain adaptation algorithms in various set-ups with different types and numbers of hand movements increasing also the available source subjects. The findings show that posture recognition can be improved by the use of prior knowledge of other intact subjects even with few available training samples. The number of training samples is reduced by an order of magnitude in domain adaptation methods to obtain the same performance achieved as when learning from scratch. Not all methods were found to perform equally well, however. The Two-Stage Weighting Framework [18] method needed a longer running time than the others and the Geodesic Flow Kernel [19] algorithm resulted in poor performance. The methods that achieved the best performance were Multi-Kernel Adaptive Learning [20] and Multi Adapt [7, 21].

III. ALGORITHMS

We, first, describe the mathematical background and the a traditional learning algorithm (III-A), then we present the domain adaptation methods used in the work (III-B).

A. Background

Let us define a training dataset $D = \{\mathbf{x}_i, y_i\}_{i=1}^N$ of N input samples $\mathbf{x}_i \in \mathcal{X}$ and corresponding labels $y_i \in \mathcal{Y}$. In the context of myoelectric classification, the inputs are the EMG signals and the labels are the movements chosen from a set of G possible classes such that the input space $\mathcal{X} \subseteq \mathbb{R}^d$ and output space $\mathcal{Y} = \{1, \dots, G\}$. The goal of a classification algorithm is to find a function $h(\mathbf{x})$ that, for any future input vector \mathbf{x} , can determine the corresponding output y . The parameters of this function are determined by solving an optimization problem over the training data. Between the algorithms that work in this manner, Support Vector Machines (SVMs) are some of the most popular.

We briefly introduce here Least-Square Support Vector Machines (LS-SVM [22]) as the starting point of the algorithms described in the next section. It is a variant of SVM and solves a classification problem (in the one-vs-all version) minimizing the squared error with an equality constraint. Given a vector \mathbf{x} with unknown label, the output hypothesis $h(\mathbf{x})$ can be written as $\langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b$. In the previous equation, \mathbf{w} and b are the parameters of the separating hyperplane, computed by the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w}, b} \left\{ \frac{1}{2} \|\mathbf{w}\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2 \right\} \\ \text{subject to: } y_i = \langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b + \xi_i, \quad \forall i \in \{1, \dots, N\}, \end{aligned} \quad (1)$$

where C is a regularization parameter and ξ denotes the prediction errors. To obtain a better solution we replaced the original training data \mathbf{x} with $\phi(\mathbf{x})$, switching from the input space to a higher dimensional feature space. Usually the function $\phi(\cdot)$ is unknown and we work directly with the kernel function $K(\mathbf{x}', \mathbf{x})$, defined as the dot product $\langle \phi(\mathbf{x}'), \phi(\mathbf{x}) \rangle$ [22]. Here we use a Gaussian kernel:

$$K(\mathbf{x}', \mathbf{x}) = e^{-\gamma \|\mathbf{x}' - \mathbf{x}\|^2} \text{ for } \gamma > 0. \quad (2)$$

B. Adaptive Learning

In the following we suppose to have K different sources, where each source is a classification model for the same set of postures. We introduce domain adaptation algorithms that construct a classification model for the same task for a new target using past experience from the sources.

1) *Multi Adapt (MA)*: This method aims to find for the target a new separating hyperplane \mathbf{w} close to a linear combination of the pre-trained source models $\hat{\mathbf{w}}^k$ [7, 21]. We solve the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w}, b} \left\{ \frac{1}{2} \|\mathbf{w} - \sum_{k=1}^K \beta^k \hat{\mathbf{w}}^k\|^2 + \frac{C}{2} \sum_{i=1}^N \xi_i^2 \right\} \\ \text{subject to: } y_i = \langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b + \xi_i. \end{aligned} \quad (3)$$

The vector $\beta = [\beta_1, \beta_2, \dots, \beta_K]^T$ with $\beta_k \geq 0$ represents the contribution of each source in the solution of the target problem and is obtained optimizing a convex upper bound to the leave-one-out misclassification loss [21]. The same paper proposes a more general case that consists in the use of different weights for different classes of the same source, e.g., $\beta_{k,g}$ is the weight associated to class g of source k . This extension is enough intuitive: it can be reasonable to assume that a user learns a movement better from one subject and another movement from a different subject. In this work we used this latter version referring to this as Multi Adapt (MA).

2) *Multi Kernel Adaptive Methods (MKAL)*: This algorithm exploits a combination of source kernels to build the new target model [20, 23]. Let us define:

$$\begin{aligned} \bar{\mathbf{w}} = [\mathbf{w}^0, \mathbf{w}^1, \dots, \mathbf{w}^K] \quad \text{and} \\ \bar{\phi}(\mathbf{x}, y) = [\phi^0(\mathbf{x}, y), \phi^1(\mathbf{x}, y), \dots, \phi^K(\mathbf{x}, y)] \end{aligned} \quad (4)$$

respectively as the concatenated vectors of the hyperplanes and the map functions in the feature space. These are both composed by $(K + 1)$ blocks: one for each source plus the block labelled with 0, referred to the original training vectors. $\mathbf{w}^k = [\mathbf{w}_1^k, \dots, \mathbf{w}_G^k]$ is the hyperplane belonging to source k , it is composed of G blocks, one for each class. The optimization problem becomes:

$$\begin{aligned} \min_{\bar{\mathbf{w}}} \left\{ \frac{\lambda}{2} \|\bar{\mathbf{w}}\|_{2,p}^2 + \frac{1}{N} \sum_{i=1}^N \xi_i \right\} \\ \text{subject to } \langle \bar{\mathbf{w}}, (\bar{\phi}(\mathbf{x}_i, y_i) - \bar{\phi}(\mathbf{x}_i, y)) \rangle \geq 1 - \xi_i, \quad \forall i, y \neq y_i. \end{aligned} \quad (5)$$

The element p sets the level of sparsity in the solution $\bar{\mathbf{w}}$ and can vary in the range $(1, 2]$. In the final prediction the

contribution coming from each source is weighted to reflect the importance of each kernel.

3) *High Level-Learning2Learn (H-L2L)*: It is an algorithm composed of two layers [24, 25]. In the first layer we build a classification model with an LS-SVM (Equation 1) for the target using a part of the original training vectors (63% from each class). We use the remaining training vectors (37% for each class) to calculate the confidence score for each source model and the target model. Thus, given a vector x , for each output class g we have $(K + 1)$ scores: one for the target $s^t(x, g)$ and K for the models of the sources $s^s(x, g)$. In the second layer, the vectors used to train the new model are the concatenation of the confidence scores of the target and the sources, obtained from 37% of the original training vectors. These are exploited to solve a multiclass LS-SVM problem with a Gaussian kernel (Equation 1).

IV. EXPERIMENTAL SETUP

The aim of this work is to tackle a situation where a new target user wants to learn to perform the same hand postures that other source subjects are already able to do. First, we need a large and public database of EMG data to make the result reliable and reproducible. The raw data are then processed to extract the most relevant characteristics. Starting from these data and the source knowledge we use domain adaptation methods to build classification models for the target. The quality of each model is estimated in terms of recognition rate. In the following we describe each step in detail.

A. Data

Data used in our work were collected in the NinaPro (Non Invasive Adaptive Prosthetics) database [9] (<http://ninapro.hevs.ch/>). According to our knowledge, it is the largest existing public database: 40 intact subjects and 11 amputees with 50 postures. Twelve *Delsys* electrodes acquired EMG data from the arm with a sampling rate of 2 *kHz*. Each subject repeated several movements 6 times, each repetition is alternated with the rest posture. To perform the exercise intact subjects used the dominant hand, while amputees mimicked with the missing arm. The postures that we take into account are 8 hand configurations and 9 movements of the wrist. Then, we are going to solve a classification problem with 18 classes: 17 movements plus the rest posture.

B. Features

A standard process is now applied to raw data [26, 27]. The initial data consists of 12 signals, one for each electrode, composed of rest-movement sequences. We separate different postures and divide each one in overlapping sliding windows (10 ms) of fixed length (200 ms). Then, we separate training and test data. The last step is feature extraction that, which determines the representation of the data for the subsequent classifier. We averaged the Mean Absolute Value, Variance and Waveform Length features based on their performance in prior studies [26] and to reduce the dependency on one specific type of representation.

C. Classifiers

We build the classification models for the sources with a non-linear SVM with Gaussian kernel, using the LIBSVM library [28]. The adaptive algorithms used to build the classification models are described in Section III-B. We also considered two baselines:

- 1) No Transfer: it uses an LS-SVM with Gaussian kernel trained on the target data only. This corresponds to learning to control the device without the help of prior knowledge.
- 2) Prior Features: it uses a linear LS-SVM trained on the prediction of source models only. This corresponds to attempting to control the prosthesis using only prior models, ignoring the target data.

In the test phase, the accuracy for each model is given by the evaluation of the agreement between real and predicted labels of the test vectors.

All the classification models are based on a non-linear SVM with Gaussian kernel. The SVM and Gaussian hyperparameters, respectively C and γ , are set with cross validation on a grid with the following values $C \in \{0.01, 0.1, 1, 10, 100, 1000\}$ and $\gamma \in \{0.01, 0.1, 1, 10, 100, 1000\}$.

D. Protocol

We ran three different experiments:

- 1) intact target subjects that exploit prior knowledge from other intact sources;
- 2) amputated target subjects that exploit prior knowledge from other amputated sources;
- 3) amputated target subjects that exploit prior knowledge from intact sources.

To create each target's model we increase a set of randomly selected training vectors in steps of 120 samples, up to a maximum of 2160. We do not balance the vectors from different classes to test the quality of each model in the worst conditions. The performance is evaluated with about $2 \cdot 10^4$ test vectors per target. Thus, we consider the trend of accuracy averaged over subjects as a function of the number of training vectors.

V. EXPERIMENTS

We evaluate the results obtained using the algorithms described in Section III-B for data from the NinaPro database (Section IV-A) processed as explained in Section IV-B. We ran three experiments. The first involves 20 random intact subjects, the second 9 no-random amputated subjects and the third both of them. In the first and second set-up, one by one, each subject represents the target problem just once, while the remaining are used as sources. In the last experiment the amputees are only the target, one by one, and the intact subjects represent only the sources. The first experiment was performed also with 10 and 30 subjects, but taking into account that the final performance slightly increases from 10 to 20 and does not change from 20 to 30, in this work we report only the analysis for 20 subjects. In the second experiment we tested all the

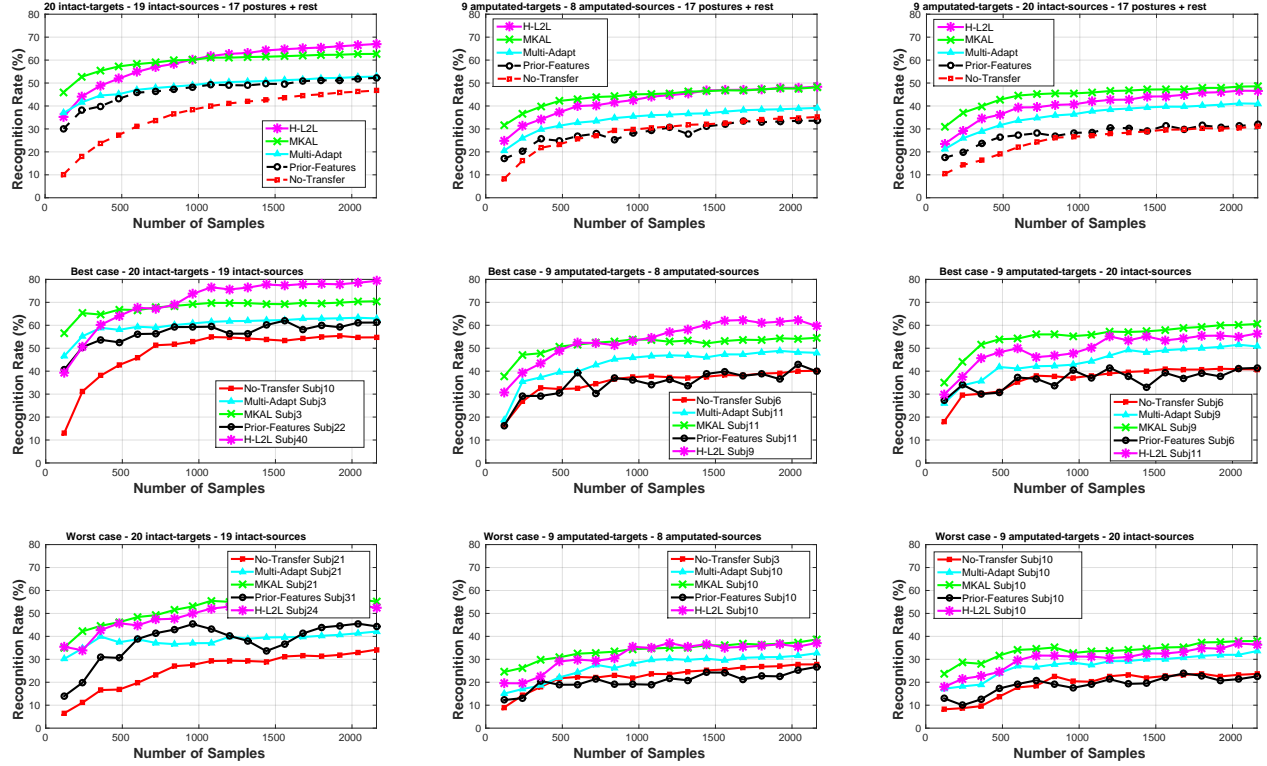


Figure 1. Classification rate obtained by averaging over all the subjects (top), for the best (middle) and for the worst cases (bottom) as a function of the number of samples in the training set: first experiment (left), second experiment (middle), third experiment (right).

11 amputated subjects, but 2 of them had only 10 electrodes because of insufficient space in the stump and they resulted lower in performance. Thus, in the following we report only the results for 9 subjects.

A. Recognition Rate Analysis

In Figure 1 we report the average, minimum and maximum performance over the set of target subjects. Despite a component of noise in the best and worst cases, the ranking between different algorithms is preserved. The domain adaptation algorithms always outperform No Transfer with an average of about: 12% (MA), 23% (MKAL) and 22% (H-L2L) in the first experiment; 6% (MA), 16% (MKAL) and 14% (H-L2L) in the second; 11% (MA), 20% (MKAL) and 15% (H-L2L) in the third. The knowledge of the sources positively affects the learning of a target, improving the performance even with few training samples. Furthermore, observe that the maximum level of performance of the No Transfer baseline is achieved by the domain adaptation algorithms using only a fraction of the target training data (i.e., from 240 up to 840 samples). In other words, prior knowledge allows us to reduce the training time by an order of magnitude, regardless of the nature of sources and target. The main difference between the three experiments concerns the absolute level of performance: the recognition rate of the first experiment is higher than the one of the second and third. This can partially be explained by the fact that classification is more difficult on EMG data from amputees (e.g., see [9]), as demonstrated

by the difference in performance of the No-Transfer baselines in the first two experiments. Furthermore, observe that the relative improvement of domain adaptation algorithms is not as large for amputees as it is for intact subjects. This is a strong indicator that there is more variability in the EMG domains of amputees and it is therefore more difficult to find an overlap between the domain of an amputee and those of other subjects.

B. Movement Classification Analysis

The recognition rate gives information about the average classification accuracy, ignoring the exact type of misclassifications by an algorithm. To investigate this recognition behaviour of each class in more depth, we show the confusion matrices in Figure 2. In the ideal case the matrix is red in the diagonal and blue outside: the predicted labels are equal to the true ones and the wrong predictions are equal to zero. We report the results with only 120 training samples to highlight the improvement when exploiting prior knowledge in the small sample setting. Focusing on No Transfer we can observe its strong dependence on the specific set of random training vectors. With only few samples we obtain a classifier that is biased toward the movements that appeared most frequently in the training data. Prior knowledge increases the recognition for all classes: even with few training samples the highest mean prediction corresponds to the true labels for the majority of movements.

It is also interesting to focus on the difference of confusion matrices of the second and third experiment for the same

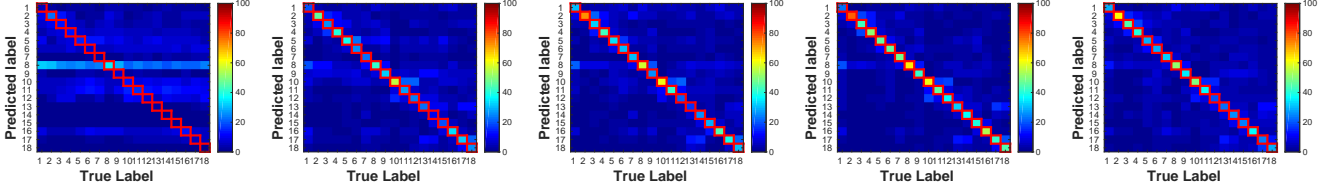


Figure 2. Confusion matrices for (from left to right) No Transfer, Prior Features, MA, MKAL and H-L2L from the first experiment with 120 training samples.

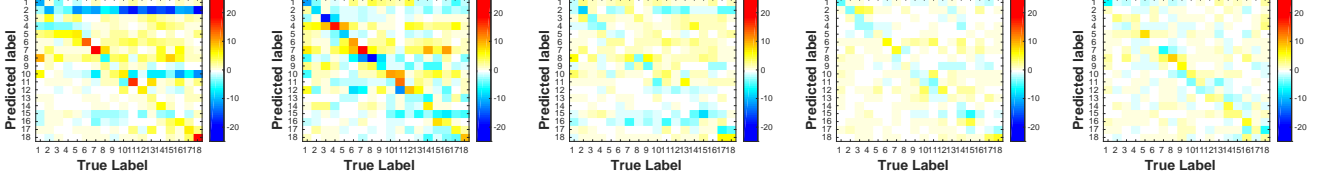


Figure 3. Difference between the confusion matrices of the second and third experiment for (from left to right) No transfer, Prior, MA, MKAL and H-L2L with 2160 training samples.

algorithm: red, blue and white colours represent respectively an increase, decrease and absence in (mis)classification (Figure 3). Looking at the diagonal there are some movements that appear differently recognized in the two experiment, but this difference varies in a range of $[-9\%, 6\%]$ for MA, $[-6\%, 6\%]$ for MKAL and $[-7\%, 8\%]$ for H-L2L exploiting all the training samples available. This highlights that an amputee does not suffer a significant learning gap exploiting intact or amputated sources. Note that the discrepancies in the No-Transfer case are caused by a different instantiation of the set of training samples. In fact, we also observe that these differences are much smaller in case of the domain adaptation algorithms, showing that using prior information has a stabilizing effect on the learning.

Focusing on single movement recognition, we want to understand the similarity in misclassifications when varying the domain adaptation algorithm or the number of training samples. For this purpose we take into account the four classes with the highest predictions for each real class (i.e., each column of the confusion matrices). We ignore the fifth prediction as it usually has a recognition lower than 5%. Comparing different algorithms for the same number of training samples, we calculate the percentage of real classes that are equal in at least 3 out of the 4 most common (mis)predictions (Table I). We did the same analysis comparing the predictions for the same algorithm with a different number of training vectors (Table II). The findings show that a class is generally misclassified with the same set of postures, regardless of the number of training samples and the algorithm used.

Finally, we want to quantify the similarity between the confusion matrices of all the experiments and classification methods to understand if a posture is learned differently when changing the target and the sources. To this end we calculate the correlations between the average recognition percentage for each class when keeping the number of training samples fixed. This correlation is computed for each combination of experiment and method, resulting in the correlation matrix

Experiment	Training Vectors	MA - MKAL	MKAL - H-L2L	MA - H-L2L
First	120	72% (13/18)	78% (14/18)	83% (15/18)
First	1080	94% (17/18)	100% (18/18)	78% (14/18)
First	2160	78% (14/18)	83% (16/18)	72% (13/18)
Second	120	61% (11/18)	67% (12/18)	33% (6/18)
Second	1080	72% (13/18)	94% (17/18)	72% (13/18)
Second	2160	72% (13/18)	94% (17/18)	56% (10/18)
Third	120	50% (9/18)	72% (13/18)	33% (6/18)
Third	1080	89% (16/18)	83% (15/18)	67% (12/18)
Third	2160	72% (13/18)	89% (16/18)	61% (11/18)

Table I
PERCENTAGE OF SIMILARITY IN THE PREDICTION OF CLASSES BETWEEN THE DIFFERENT ADAPTIVE ALGORITHMS FOR THE SAME NUMBER OF TRAINING VECTORS.

Experiment	Training Vectors	MA	MKAL	H-L2L
First	120-1080	94% (17/18)	89% (16/18)	78% (14/18)
First	1080-2160	100% (18/18)	94% (17/18)	94% (17/18)
Second	120-1080	33% (6/18)	72% (13/18)	61% (11/18)
Second	1080-2160	94% (17/18)	100% (18/18)	94% (17/18)
Third	120-1080	39% (7/18)	83% (15/18)	72% (13/18)
Third	1080-2160	78% (14/18)	100% (18/18)	89% (16/18)

Table II
PERCENTAGE OF SIMILARITY IN THE PREDICTION OF CLASSES WHEN THE NUMBER OF TRAINING VECTORS CHANGES FOR THE SAME METHODS.

shown in Figure 4. The two diagonal blocks of the matrix contain the correlations between intact targets (first experiment) and amputated targets (second and third experiments). The blocks outside the diagonal hold the correlation between intact and amputated targets. These findings show that the correlation between the second and third experiment is high, thus an amputee learns a movement with the same simplicity or difficulty from intact and amputated sources. Instead the correlation between amputated and intact targets is low, meaning that the set of movements that can be learned with relative ease is not identical in the two cases.

VI. CONCLUSION AND COMPARISON

In this work we have dealt with the control of non-invasive myoelectric prostheses. We have investigated the variation

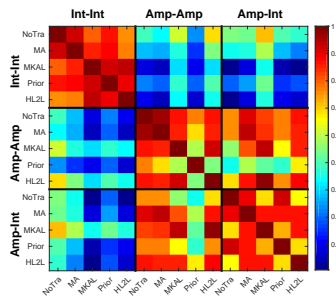


Figure 4. Correlation matrix for each algorithm and experimental setting. II, AA and AI stand respectively for: intact-intact (i.e., first experiment), amputees-amputees (i.e., second experiment) and amputees-intact (i.e., third experiment).

in learning performance when prior information from other subjects is used and whether the type of knowledge (from amputees or intact) is significant for an amputee. Our findings show that prior knowledge positively affects the trend of the recognition curves, increasing the performance and reducing the training time by an order of magnitude. The main result concerns the comparison when using either intact or other amputated subjects as sources for an amputated target. The obtained mean performance appears unchanged for all the algorithms, meaning that an amputee learns with the same effectiveness from intact or amputated sources. The same conclusion emerges from the correlation matrices that show a high agreement between both settings: an amputee learns a movement with the same simplicity or difficulty from intact and amputated subjects. This result is important since there is a large number of potential intact source subjects with respect to amputees, for who participating in data collection is difficult or even painful.

REFERENCES

- [1] D. J. Atkins, D. C. Y. Heard, and W. H. Donovan, "Epidemiologic overview of individuals with upper-limb loss and their reported research priorities," *Journal Of Prosthetics And Orthotics*, vol. 8, no. 1, pp. 2–11, 1996.
- [2] M. Zecca, S. Micera, M. C. Carrozza, and P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," *Critical Reviews in Biomedical Engineering*, vol. 30, no. 4–6, pp. 459–485, 2002.
- [3] S. Micera, J. Carpaneto, and S. Raspopovic, "Control of hand prostheses using peripheral information," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 48–68, Octobre 2010.
- [4] B. Peerdeman, D. Boere, H. Witteveen, R. H. in 't Veld, H. Hermens, S. Stramigioli, H. Rietman, P. Veltink, and S. Misra, "Myoelectric forearm prostheses: State of the art from a user-centered perspective," *Journal of Rehabilitation Research and Development*, vol. 48, no. 6, pp. 719–738, 2011.
- [5] C. Castellini and P. P. van der Smagt, "Surface EMG in advanced hand prosthetics," *Biological Cybernetics*, vol. 100, no. 1, pp. 35–47, 2009.
- [6] D. Farina, C. Cescon, and R. Merletti, "Influence of anatomical, physical, and detection-system parameters on surface EMG," *Biological Cybernetics*, vol. 86, no. 6, pp. 445–456, 2002.
- [7] T. Tommasi, F. Orabona, C. Castellini, and B. Caputo, "Improving control of dexterous hand prostheses using adaptive learning," *IEEE Transactions on Robotics*, vol. 29, no. 1, pp. 207–219, 2013.
- [8] N. Patricia, T. Tommasi, and B. Caputo, "Multi-source adaptive learning for fast control of prosthetics hand," in *22nd International Conference on Pattern Recognition*, 2014, pp. 2769–2774.
- [9] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatsidis, F. Bassetto, and H. Müller, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Scientific data*, vol. 1, 2014.
- [10] M. Atzori, A. Gijsberts, S. Heynen, A.-G. M. Hager, O. Deriaz, P. Van Der Smagt, C. Castellini, B. Caputo, and H. Müller, "Building the ninapro database: A resource for the biorobotics community," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechanics (BioRob)*. IEEE, 2012, pp. 1258–1265.
- [11] M. Atzori, A. Gijsberts, I. Kuzborskij, S. Heynen, A.-G. Mittaz Hager, O. Deriaz, C. Castellini, H. Müller, and B. Caputo, "Characterization of a benchmark database for myoelectric movement classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2014.
- [12] D. Graupe and W. K. Cline, "Functional separation of EMG signals via ARMA identification methods for prosthesis control purposes," *IEEE Transactions on Systems, Man and Cybernetics*, no. 2, pp. 252–259, 1975.
- [13] C. Castellini, A. E. Fiorilla, and G. Sandini, "Multi-subject / daily-life activity EMG-based control of mechanical hands," *Journal of Neuroengineering and Rehabilitation*, vol. 6, no. 41, 2009.
- [14] T. Matsubara, S. Hyon, and J. Morimoto, "Learning and adaptation of a stylistic myoelectric interface: EMG-based robotic control with individual user differences," in *ROBIO*. IEEE, 2011, pp. 390–395.
- [15] J. W. Sensinger, B. A. Lock, and T. A. Kuiken, "Adaptive pattern recognition of myoelectric signals: Exploration of conceptual framework and practical algorithms," *IEEE Transactions on Neural systems and rehabilitation engineering*, vol. 17, no. 3, 2009.
- [16] F. Orabona, C. Castellini, B. Caputo, A. E. Fiorilla, and G. Sandini, "Model adaptation with least-square svm for adaptive hand prosthetics," in *IEEE International conference on Robotics and Automation*, 2009.
- [17] R. Chattopadhyay, N. Krishnan, and S. Panchanathan, *Topology preserving domain adaptation for addressing subject based variability in SEMG signal*, 2011, vol. SS-11-04, pp. 4–9.
- [18] Q. Sun, R. Chattopadhyay, S. Panchanathan, and J. Ye, *A two-stage weighting framework for multi-source domain adaptation*, 2011.
- [19] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2066–2073.
- [20] F. Orabona, J. Luo, and B. Caputo, "Online-batch strongly convex multi kernel learning," in *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, 2010, pp. 787–794.
- [21] T. Tommasi, F. Orabona, and B. Caputo, "Learning categories from few examples with multi model knowledge transfer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 5, pp. 928–941, 2014.
- [22] J. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. World Scientific, 2002.
- [23] F. Orabona, J. Luo, and B. Caputo, "Multi kernel learning with online-batch optimization," *Journal of Machine Learning Research*, vol. 13, pp. 227–253, 2012.
- [24] T. Tommasi, F. Orabona, and B. Caputo, "Discriminative cue integration for medical image annotation," *Pattern Recognition Letters*, vol. 29, no. 15, pp. 1996–2002, 2008.
- [25] N. Patricia and B. Caputo, "Learning to learn, from transfer learning to domain adaptation: A unifying perspective," in *2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014*. IEEE Computer Society, 2014, pp. 1442–1449.
- [26] I. Kuzborskij, A. Gijsberts, and B. Caputo, "On the challenge of classifying 52 hand movements from surface electromyography," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2012, pp. 4931–4937.
- [27] A. Gijsberts, M. Atzori, C. Castellini, H. Müller, and B. Caputo, "Movement error rate for evaluation of machine learning methods for sEMG-based hand movement classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 735–744, 2014.
- [28] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.